Draft Motion Imagery Quality Equation (MIQE)

Dr. Darrell L. Young & Dr. Tariq Bakir
Motion Imagery Quality Metrics
Contractor
youngdl@nga.mil
703 262 4418
tbakir@harris.com
321-984-6649

March 2009
JACIE
UNCLASSIFIED

Draft Motion Imagery Quality Equation (MIQE)

- **Purpose:**
  - **Provides a method to predict Video National Imagery Interpretability Rating Standard (V-NIIRS) given system technical (mission + optical) parameters.**
  - **Provides a method to predict V-NIIRS, given existing imagery and metadata.**
  - **Provides a method to compute probability of task success so that motion imagery quality can be included in fusion and dependence chains**
- **Beta-MIQE is NOT approved for mission planning, procurement specification, or any other use. It is provided for comment only.**
- **Bottom Line:**
  - Beta-MIQE provides a method to convert technical parameters into V-NIIRS equivalents which are more easily used by analysts. Supports problem driven collection, and retrieval.

Components of Object Interpretability

- ***Detection***

Is the perceptibility of an object's (which may be a target image) presence at a particular location, distinguishable from its surroundings.

- ***Classification***

Is the determination of whether a detected object is a member of a particular set of possible targets or non-targets (e.g., wheeled versus tracked vehicles).

- ***Recognition***

Is the determination that a target belongs to a particular functional category (e.g., a tank, a truck, an armored personnel carrier, etc.).

- ***Identification***

Is the most detailed level of discrimination of particular relevance for military target acquisition, as discussed shortly (e.g., a T-72, T-62, M1, or M60 tank).

Intelligence Interpretability vs. Visual Preference.

- High Mean Opinion Score (MOS), Low V-NIIRS example:
  - Lightly compressed, low resolution motion imagery can be pleasing to the eye, but impossible for fulfillment of interpretability task requirement.

- Low MOS, HIGH V-NIIRS example
  - Heavily compressed, high resolution motion imagery can be annoying, but meet interpretability thresholds.

- **Consumer preference as measured by MOS does not map to intelligence interpretability**

V-NIIRS Defines Object AND Activity Recognition

Each of the written criteria contains specific components separated by a bullet point to add clarity and aid readers understanding of the content.

- 9 levels of quality
- 7 orders of battle
  - Aircraft
  - Electronic
  - Ground
  - Missile
  - Naval
  - Cultural
  - Human

- Analyst Task
- Object of Interest
- Associated Activity or Behavior
- Environment
- (Object Reference Examples)

V-NIIRS Level 3

Visually track movement of ● an identified Heavy Cargo/Passenger Aircraft ● during taxi or tow ● at a primary airfield/airport installation.● (Aircraft Length: 150ft or more, eg. MD11, A300, B747, B767, DC8)
Visually track the movement of ● Unidentified radar/radar support vehicles ● in column/convoy or deploying ● in the vicinity of a known EW or SAM radar site ● (4 to 8 vehicles with total column length 150ft plus)
Visually track the movement of ● an unidentified military convoy of company size or larger, possible armor or mechanized infantry● in a column or "road march"● on an open highway/roadway● (4 to 8 vehicles with total column length 150ft plus)
Visually track the movement of ● Convoy of intermediate-range ballistic missile (IRBM) transporter and support vehicles ● during deployment or road march ● on an improved road near missile base, launch site or silo ● (Dong Feng 4, Taepodong 2, Agni 3/4, Shaheen 2/3: transporter with support vans - convoy length 60m or more)
Visually track the movement of ● an unidentified coastal patrol craft ● conducting normal operations ● at sea several miles beyond a harbor or port ● (Example, US Cyclone Class: average 175ft length, 25ft beam)
Visually track the movement of ● an unidentified tractor-trailer rig convoy of 3 or more vehicles ● driving in a column formation ● on the open highway ● (big-rig tractors with long trailers, total convoy length 150ft or more)
Visually track the movement of ● an unidentified convoy of 3 or more sea/land containers ● driving in a column formation ● exiting a railyard or port facility ● (big-rig tractors hauling Sea/Land Containers total convoy length 150ft or more)

V-NIIRS 11

Visually confirm the movement of ● the fingers/hand of a ground crew/mechanic changing the socket on a ratchet/socket wrench ● while servicing any aircraft or support vehicle ● at any airfield, base, or aircraft maintenance facility ● (Socket able to fit in palm of workers hand)
Visually confirm movement of ● an individual's mouth/jaw ● while speaking into a bluetooth wireless mobile phone earpiece ● in an crowded public area or pedestrian walkway ● (Average sized person wearing an over the ear device with internal or boom microphone, avg diameter: 1 to 2 inch)
Visually confirm the movement of ● the fingers and hands of an individual holding a fragmentation grenade ● as the weapon's safety is released and the the device is readied ● at a practice range, during live fire exercise, or during an engagement ● (spherical or cylindrical device, palm sized with metal pull ring/pin and spring loaded spoon: 2.5in - 3in diameter)
Visually confirm the movement of ● individual's fingers and hands while aiming a shoulder fired anti tank missile ● as they release safetys and arm the device ● at a tactical position in a rural or urban environment ● (Individuals of average height and weight holding AT-4 or RPG)
Visually confirm the movement of ● an individual combat swimmer's hands and fingers ● as they check out and test scuba equipment ● on a light surface combatant, i.e., patrol boat near the littoral zone ● (individual of average size and weight)
Visually confirm the movement of ● an individual pedestrian's hands and fingers ● as they make change or sort coins ● in a busy open market or square ● (individual of average height & weight, sorting coins in a change purse or the palm of one hand)
Visually confirm the movement ● of an individuals hands and fingers ● as they communicate through sign language ● in an open public area ● (i.e., fully interpret sign language to include the spelling out of individual letters)
Visually isolate and investigate an individual or group based on the movement of ● their hands, fingers, and face ● as they observe the movement of a protected individual ● from a crowd, behind a rope line or police barrier ● (i.e., isolate behavior to determine if an individual or individuals pose an immediate threat to a VIP)
Visually isolate and investigate an individual apparently burdened with significant concealed weight, based on their gait ● their posture, hands, fingers, and overall body language ● as they meander into a crowd ● in a public square, market, or shopping mall ● (i.e., based on gait, determine if an individual is a public threat, concealing an explosive vest or belt)
Visually isolate suspicious movement/behavior of ● the hands and fingers of a suspect individual or known operative ● as they leave an inconspicuous signal or message ● on a lightpole or mailbox in a crowded urban street ● (i.e., traditional espionage tradecraft: a chalkmark or sticker on a predetermined location to signal a meeting or dead-drop)

Beta MIQE

$$MIQE = M - a\log10(GSD_{GM}) + 2*\log10(Q) + b\log10(RER_{GM}) - (0.656)(H) - (0.344)(G/\min(SNR, \beta C)) \quad (27)$$

M=11.6, a=3.32 and b=1.559 for RER>=0.9 and

M= 11.53, a=3.16 and b=2.817 for , RER < 0.9.

Parameter	Minimum	Maximum
GSD	0.75 cm	220 cm
RER	.2	1.3
Overshoot, H	.9	1.9
Noise Gain, G	1	19
SNR	2	130
β, SNR-to-Contrast	100	130
Peak Foreground Discontinuity, Δ_T	0	1.3
Horizontal Trend, T_{TH}	0	TBR
Horizontal Jitter, σ_{TH}	0	TBR
Vertical Trend, T_{TV}	0	TBR
Vertical Jitter, σ_{TV}	0	TBR
Elevation angle, ϕ	30 degrees	90 degrees
GSS parameter, K	1	-
Modulation Contrast	0.15	1.0
Q	1	2

Review: MTF to the Edge

(and back)

convolutions

$LSF_{sys}=LSF_1 \circledcirc LSF_i \circledcirc LSF_N$

$$ER_x(x) = 0.5 + \frac{1}{\pi}\int_0^\infty \frac{MTF_x(f_x)}{f_x}\sin(2\pi f_x x)df_x$$

products

$MTF_{sys}=\prod_{i}^{N} MTF_i$

differentiate

Fourier Transform

Magnitude

RER_{sys} ↔ Edge

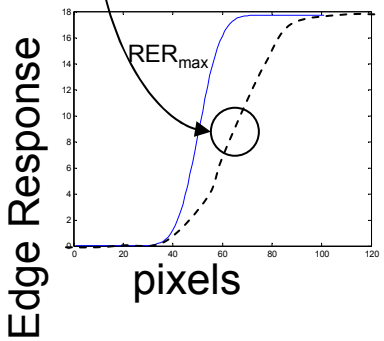
Line **S**pread **F**unction

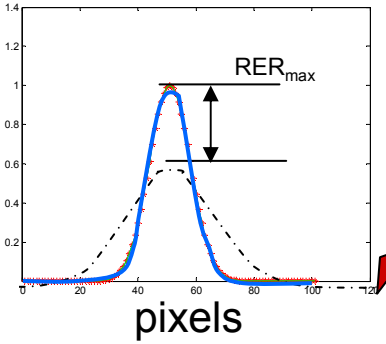
Modulation **T**ransfer **F**unction

integrate

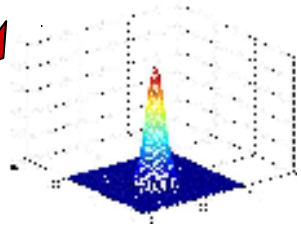
Inverse Fourier Transform

slope



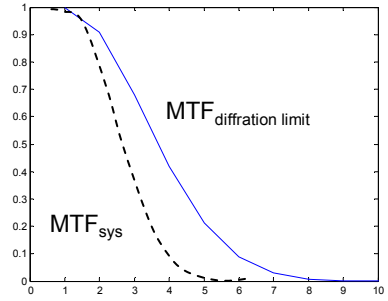


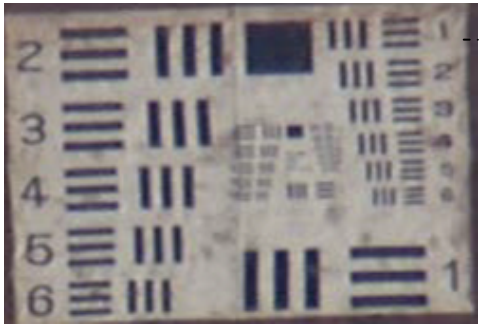
1-D slice of PSF



Point Spread Function (PSF)

Contrast Modulation





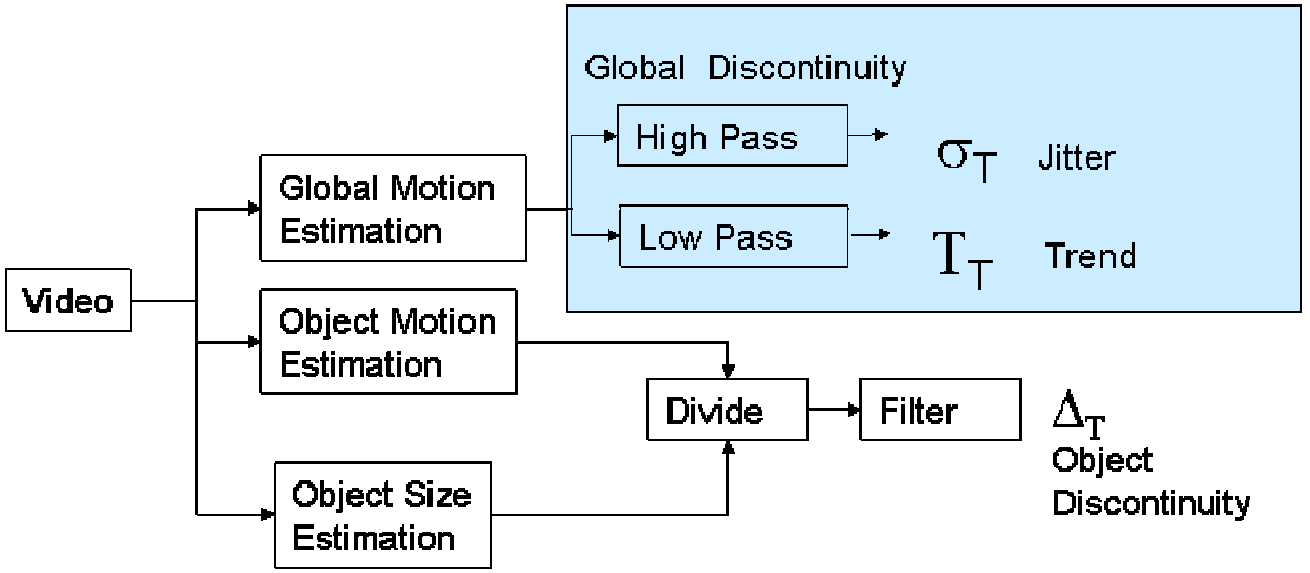
Strehl Ratio~ $\frac{RER_{sys}}{RER_{max}}$

Motion can degrade overall system MTF for multiple reasons:

Intra-Frame	Inter-Frame	Human Eye
$\sigma_{\tau,}$ Jitter smear	$\sigma_{T,}$ Jitter motion	DVA reduced
$T_{\tau,}$ Trend smear	$T_{T,}$ Trend motion	DCSF reduced
$\Delta\tau,$ Target motion smear	$\Delta_{T,}$ Target Motion	

Interframe Motion Factors Affecting Interpretability

- Camera Motion (Global)
  - Random global motion (jitter, σ_T)
  - Trend global motion (panning and tracking, T_T)
- Target Motion
  - Motion-based temporal aliasing (object discontinuity, Δ_T)

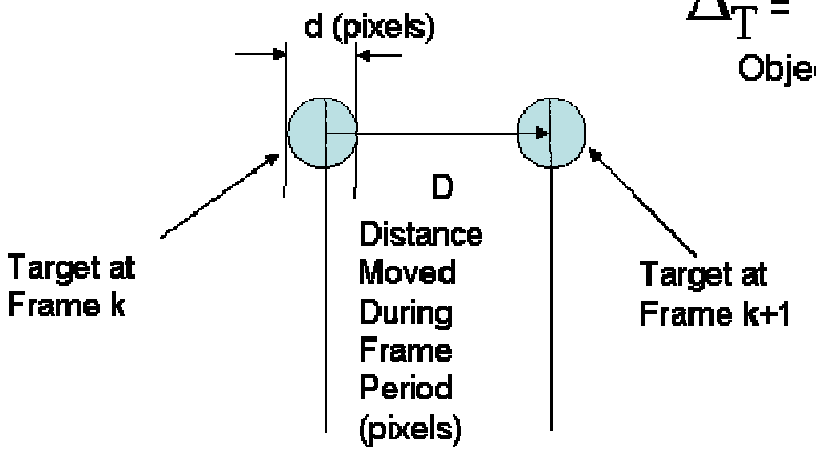


Examples of Spatial and Temporal Resolution

MIIRS	Object	Spatial Resolution (cm)	Action	Event Duration (sec)	Minimum Sampling Rate (FPS)
3	semi in convoy	200.0	making turn	10.0	1.0
4	bus in light traffic	100.0	making turn	5.0	2.0
5	lone car	50.0	making turn	3.0	3.3
6	car in traffic	25.0	changing lanes	2.0	5.0
7	motorcycle	12.5	changing lanes	1.0	10.0
8	people	6.3	getting into car	1.0	10.0
9	sub-groups in crowd	3.1	movements	1.0	10.0
10	the body and limbs of participants	1.6	confirm a conversation is underway based on the movement	0.5	20.0
11	an individual's mouth/jaw	0.8	while speaking into a cell phone	0.3	30.0

Δ_T ,Discontinuity

- Object Discontinuity (Δ_T) : is the ratio of target motion per-frame (D) to target size (d).



$$\Delta_T = \frac{\text{Object movement in Pixels}}{\text{Object size in direction of motion}}$$

Discontinuity represents the temporal effects to quality in motion imagery, or full motion video.

Practical rationale for alignment of the NIIRS, and V-NIIRS scales

- Huge cross-training and cost savings benefit. NIIRS is already well-known and accepted across IC/DoD and allied communities.

- The spatial alignment of NIIRS, and V-NIIRS enables use of the GIQE for the spatial resolution aspect of motion imagery.

- The temporal aspect is addressed by setting thresholds on discontinuity, that result in derived requirements on framerate, and stability.

Suggested Field-of-View

Implies requirement for HD

V-NIIRS Level	GSD (approx)	GSD (approx)	approximate object length	FOV	RES	Format	object length to FOV factor
	meters	**inches**	meters	meters (on long side)	min. # of pixels (on long side)		
11	0.008	0.30	0.2	10	1292	HD	42
10	0.015	0.60	0.5	18	1169	HD	38
9	0.030	1.20	0.9	32	1046	HD	34
8	0.061	2.40	1.9	56	923	HD	30
7	0.122	4.80	3.8	98	800	HD	26
6	0.244	9.60	7.5	165	677	ED	22
5	0.488	19.20	15.0	270	554	ED	18
4	0.975	38.40	30.0	420	431	ED	14
3	1.951	76.80	60.0	600	308	ED	10

Comparison to Army Target Acquisition Model

The critical dimension of the target is defined to be the square root of the width, height product.

$$d_c = \sqrt{W_{tgt} H_{tgt}} \,. \tag{32}$$

The original Target Acquisition Model (TAM) empirically measured the number cycles needed for 50% probability of successful completion of task. More difficult tasks require more cycles as shown

Empirical Number of Cycles (N50) Across Critical Dimension

Task	Description	2-D cycles (N50)
Detection	Reasonable probability that blob is a military vehicle	0.75
Recognition	Class discrimination (truck, tank, etc.)	3.0
Identification	Object discrimination (M1A, T-62, or T-72)	6.0

$$P(N) = \frac{\left(N/N_{50}\right)^{2.7+0.7\left(N/N_{50}\right)}}{1+\left(N/N_{50}\right)^{2.7+0.7\left(N/N_{50}\right)}}$$

The probability for a given number of cycles N across a target is determined by,

Two-handed hand held objects example

Application of the original TAM model gives a probability of correct identification of
V-NIIRS 9 = 0.90 ASSUMING GOOD RER and SNR!
V-NIIRS 8 = 0.73

REFERENCE: Steve Moyer, Eric Flug, Timothy C. Edwards, Keith Krapels, John Scarbrough, "Identification of handheld objects for electro-optic/FLIR applications", Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XV, edited by Gerald C. Holst, Proc. of SPIE Vol. 5407